A Study of Exploring Different Schedules of Spacing and Retrieval Interval on Mathematics Skills in ITS Environment

Xiaolu Xiong, Joseph E. Beck

Department of Computer Science
Worcester Polytechnic Institute
100 Institute Road, Worcester, MA

xxiong@wpi.edu, josephbeck@wpi.edu

Abstract. The present study was designed to help answer several questions regarding the impact of spacing and expanding retrieval practice on mathematics skills. For this study, we set up four different interval schedules (1 day; 4 days; 7 days; 14 days) in an ITS environment, and examined the impact on retention performance by comparing results across groups. There were significant performance differences on different groups of students, and all fours groups of students showed small declines in the retention performance with longer intervals. Furthermore, we examined students with high-, medium-, and low-knowledge of skills, and found a strong effect on retention performance with the basis of initial performance on skills. In addition, students with weaker knowledge showed a much more rapid forgetting than students with higher knowledge. These results suggest retention intervals should probably not be fixed, but should vary based on the student's knowledge of the skill.

Keywords: knowledge retention, retrieval practice, spacing effect, intelligent tutoring system

1 Introduction

Expanding retrieval practice is based on the robust memory phenomenon known as the spacing effect, in which memory for repeated items is better when repetitions spaced apart rather than massed together [5, 6]. In expanded retrieval, these repetitions are spaced increasing intervals, making it necessary to retain the skill for longer and longer amounts of time before one attempt to retrieve it. This effect is specifically important to a cumulative subject as mathematics: we are more concerned with students' capability to remember the knowledge that they acquired over a long period of time.

adfa, p. 1, 2011.

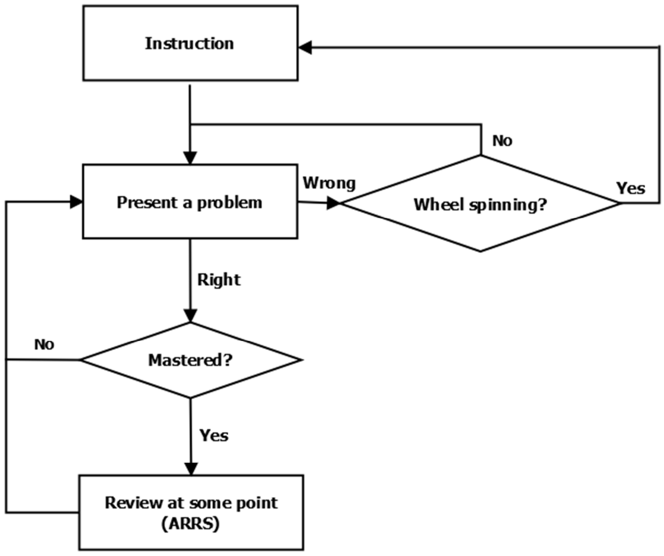

Fig. 1. The enhanced ITS mastery learning cycle

Inspired by the notion of robust learning [2] and the design of the enhanced ITS mastery cycle (Figure 1) proposed by Wang and Beck [7], we developed and deployed a system called the Automatic Reassessment and Relearning System (ARRS) to make decisions about when to review skills which students have mastered. ARRS is an implementation of expanding retrieval in the ITS environment, Unlike most ITS system [4] which the tutoring stopped if the student mastered a given skill, ARRS assumes that if a student masters a skill with three correct responses in a row, such mastery is not necessarily an indication of long-term retention. Therefore, ARRS will present the student with a reassessment test on the same skill at expanding intervals spread at least 3 months of schedule, that is firstly 7 days after mastery, then 14 days, 28 days and 56 days after the previous test. If a student fails the reassessment test, ASSISTments will give him an opportunity to relearn the skill.

We refer to the number of problems required to achieve mastery as the *mastery speed*, it represents a combination of how well the student knew this skill originally, and how quickly he can learn the skill. We observed that, in general, the slower the mastery speed, the lower the probability that the student can answer the problems in the retention test correctly. Students who mastered a skill in 3 or 4 problems had an 82% chance of responding correctly on the first retention test, while students who took over 7 attempts to master a skill only had a 62%.

Previous studies showed that mastery speed is an extremely important feature for predicting student's retention performance and has a long term effect on students' retention performance [8]. According to these results, we can say that students with different mastery speed have different retention patterns, so we decided to start the exporting the optimal retrieval schedules for different levels of students.

2 An experiment on different schedules of retention interval

We first conducted an experiment to investigate how different retention intervals affect student retention performance. There were several objectives for this experiment. A central goal was to investigate knowledge-related differences in terms of spacing and retention interval. As we mentioned before, students who receive retention tests have demonstrated mastery in the initial problem set, which we refer to as the mastery learning problem set. We already observed these students have significantly differences in the fixed-schedule retention tests. Thus, it is worth to find out how mastery speed affects the retention performance given different intervals. This experiment tested students with different retention intervals to explore this question.

The participants were 672 middle and high school students from 34 classes. Teachers of these classes enabled ARRS in ASSISTments voluntarily, and they assigned mathematics mastery learning problem sets according to whatever instructional content they would normally cover in class. Teachers also required their students to use ASSISTments to finish their homework on a daily basis. Students were randomly allocated to one of four conditions which applied with different retention intervals: 174 students were assigned to the 1-day condition, 170 students were assigned to 4-day retention test condition, 162 student and 166 students were assigned to 7-day and 14-day condition. Students worked on their assignments in various environments include school computer labs, home computers and mobile devices. Prior to this experiment, students and teachers already had experiences of using ASSISTments and working with ARRS.

Students were randomly assigned to one of four retention interval conditions: 1-day, 4-day, 7-day, or 14-day. The differences among these conditions were the interval between achieving mastery and receiving the reassessment test. For example, Students in the 1-day condition received the corresponding retention tests the day after they finished the mastery learning problem sets; while students in 14-day conditions received reassessment test 14 days after they finished the mastery learning problem sets. It is important to notice that all reassessment tests were released only on weekdays; this particular behavior of ARRS was designed to cooperate with teachers, and it delayed the assigning of the retention tests which were scheduled to be released on Saturdays and Sundays.

This experiment began on September 15, 2013 and ended on December 15, 2013. During these three months, students constantly received mastery learning problem sets as homework assignments from their teachers. Once they answered three consecutive questions correctly in a mastery learning problem set, a retention test was scheduled based on which condition a student was in and ready to be assigned (e.g., 1, 4, 7, or 14 days after mastery). For mastery learning problems sets, to finish on time, students were required to complete it within one day of when the teacher assigned it. Similarly, for ARRS tests, which were generated by ASSISTments according to the appropriate schedule interval, students had one day to complete these tests. However, it was not uncommon for students to not always complete assignments on time.

3 Results and Discussion

In this study, we asked whether a different retention interval would affect students' retention performance. We were particularly interested in whether or not longer spacing would impede students' retention. In order to determine if different retention interval affected students' performance, we examined students' retention test performance in different conditions.

As we expected, students in longer retention interval had lower retention performance than students in shorter retention interval, but none of the differences are particularly large, even the 1-day performance (80.4%) and 14-day performance (76.0%) only differed by 4.4%. We also noticed that students in the 4 days and 7 days conditions had very close retention performance, namely 77.6% and 77.5%, and this can be explained by the some portion of 4 days retention tests had been delayed one or two days to skip weekends.

When considering whether there were changes in retention performance of students with different mastery speed, we grouped the data by three identified mastery speed bins, then we also examined students' retention test performance. Table 1 shows the retention performance by mastery speed and retention interval.

Table 1. Retention performance by mastery speed and retention interval

	All retention tests (maximizes external validity)		Retention tests completed on time (maximizes internal validity)	
Retention test delay	# tests	% correctness	# tests	% correctness
mastery speed 3 - 4				
1 day	1186	84.4%	462	85.1%
4 days	1169	82.2%	389	84.6%
7 days	1171	81.7%	409	84.1%
14 days	1233	81.2%	419	83.8%
mastery speed 5 - 7				
1 day	467	77.9%	184	75.5%
4 days	432	76.2%	149	73.2%
7 days	362	77.1%	147	72.9%
14 days	420	73.1%	150	72.7%
mastery speed > 7				
1 day	280	67.5%	110	70.0%
4 days	320	62.8%	111	65.8%
7 days	267	59.6%	105	68.6%
14 days	243	54.8%	85	60.0%

The left part of Table 1 shows how students performed on retention tests, and includes data from all students. Including data from all students' results in high external validity as it ensures that our results generalize to other, similar, populations of learners. However, we have seen some tests were completed more than one week later after they were due. Including such data in the study makes it difficult to determine which experimental condition the student was in. How should we analyze students who were in the 7-day condition but completed their retention test 14 days later?

To account for students not being conscientious in completing retention tests on time, we have selected tests which were finished on time (finished no more than one day after released and made available to students). As a result, performance on these tests reflects retention performance on the intervals specified by the study. That is, a student in the 7-day condition was answering his retention test after a delay of between 7 and 8 days, but 14 days would not be possible. Although this approach maximizes internal validity, it also introduces a selection bias. Students who finish their assignments on time are not a random sample of the population, but rather are those who watch their assignment schedules more closely, and those who cared more about finishing assignments on time. These non-random selection effects make these students not perfectly representative of the population as a whole. This tension between internal and external validity is common in field research, and we present both sets of data.

In all students, we have seen consistent decrease in retention performance with longer retention intervals, whether they were high mastery level, medium mastery level or low mastery level students. The results from Table 1 also demonstrated a main effect of mastery speed on retention performance: students with slower mastery speed had significantly lower performance than students with a faster mastery speed ($p \approx 2.2 \times 10^{-27}$); this statement is true even when we comparing 1-day performance of students with slow mastery speed versus 14-day performance of students with fast mastery speed (67.5% for mastery speed > 7 versus 81.2% for mastery speed on 3 or 4). A large and interesting effect is that students with slower mastery speed had larger decrease in retention performance as retention intervals got longer. This interaction effect was statistically reliable ($p \approx 3.4 \times 10^{-22}$). For example, high mastery level student had a decrease of 3.2% between 1 day tests and 14 days tests but retention performance of low mastery level students dropped 12.7%. The horizontal comparisons on Table 1 also suggest that students who finished test on scheduled intervals were more likely to retain skills, confirming our suspicion above about these students not being a representative sample.

4 Contributions, Future work and Conclusions

As this paper contributes to a large body of literature empirically demonstrating the effects of spaced learning, it makes three unique contributions. First, this paper studied actual effects of spaced learning over long time period for mathematics materials and practices whereas most ITS studies were focused on shorter term and only few looked effects over time. Second, this experiment investigated the concept of finding the optimal retention interval by using mastery speed for students with different mastery speed. Moreover, this study suggested the necessity of retention tests as a measurement method of robust learning.

Our goal is to find the optimal spacing schedules for students and the best way to boost their performance in long-term mathematics learning; there are so many open problems worth of future research: Is there a better to predict who will retain a skill? Do these mistakes indicate lack of effort or interest on the student's part, or a genuine lack of knowledge? What should we do after students fail a retention test, should we just reply on the connection between well-learned procedural skills and long-term retention [1]? We are also interested in interventions that can decrease the rate of wheel spinning [3]. Most importantly, there are some very challenging problems that we believe can be answered in our following studies. First, do assigning high frequent retention tests and relearning assignments to low knowledge student help to improve their mastery level? And what other tutoring methods we can use if a student fails to retain a skill?

This paper presents the first study of exploring the optimal spacing schedule in learning mathematics skills. With the experiment data we collected, we revealed the relationships between master speed and retention performance in different retention intervals, and most importantly, these relationships will help dictate which learning schedules and memory techniques are most suitable for learning and retrieving.

5 Acknowledgments

We want to acknowledge the funding on NSF grant DRL-1109483 as well as funding of ASSISTments. See here (http://www.webcitation.org/67MTL3EIs) for the funding sources for ASSISTments.

6 References

1. Anderson, J. R.: Rules of the mind. Routledge (1993)
2. Baker, R.S.J.d., Gowda, S., Corbett, A., Ocumpaugh, J.: Towards Automatically Detecting Whether Student Learning is Shallow. In Proceedings of the International Conference on Intelligent Tutoring Systems, 444-453 (2012)
3. Beck, J. E., & Gong, Y.; Wheel-Spinning: Students Who Fail to Master a Skill. In Artificial Intelligence in Education (pp. 431-440). Springer Berlin Heidelberg (2013)
4. Beck, J.E., et al.: Predicting student help-request behavior in an intelligent tutor for reading. The 9th International Conference on User Modeling (2003)
5. Crowder, R.: The effects of repetition. In Principles of learning and memory, pp. 264–321. Hillsdale, NJ: Lawrence Erlbaum Associates (1976)
6. Melton, A. W.: Repetition and retrieval from memory. Science, 158,532. (1967)
7. Wang, Y., Beck, J.: Using Student Modeling to Estimate Student Knowledge Retention. In Proceedings of the 5th International Conference on Educational Data Mining, 201-203 (2012)
8. Xiong, X., Li, S., Beck, J.: Will you get it right next week: Predict delayed performance in enhanced ITS mastery cycle. The 26th International FLAIRS Conference (2013)